AD-A250 543

REPORT DOCUMENTATION PAGE

Form Approved
OPM No. 0704-0188

2

1. AGENCY USE ONLY *(Leave Blank)*

2. REPORT DATE
1989

3. REPORT TYPE AND DATES COVERED
Unknown

4. TITLE AND SUBTITLE
Statistical Inference and Causal Reasoning

5. FUNDING NUMBERS
DAAB10-86-C-0567

6. AUTHOR(S)
Jay C. Weber

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Department of Computer Science
University of Rochester
Rochester , NY 14627

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army CECOM Signals Warfare Directorate
Vint Hill Farms Station
Warrenton, VA 22186-5100

10. SPONSORING/MONITORING AGENCY REPORT NUMBER
92-TRF-0028

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Statement A, Approved for public release, distribution unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*
In this paper, we show how degrees of belief about causal predictions can be derived from statists about the truth of properties over time. By using statistical information analogous to traditional non-statistical rules of causation and knowledge abpit a specific time point, predictions can be made about both the change and persistence of properties for the next time point with some degree of belief. We show how to incrementally compute this degree of belief by combining statistics conditioned on successively larger subsets of the reasoner's knowledge. Furthermore, we solve the qualification problem through a powerful heuristic that builds these subsets by considering properties with highest impact first. This heuristic ignores relatively unlikely, redundant, or unrelated properties when deriving a prediction, while directing the focus along causal chains. The iterative formula and this heuristic define an algorithm that prduces predictions with quickly increasing confidence, allowing computational resources to trade off against accuracy.

14. SUBJECT TERMS
algorithms, statistical interference, causal reasoning, artificial intelligence, data fusion, probabilistic reasoning

15. NUMBER OF PAGES
10

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UL

Statistical Inference and Causal Reasoning

Jay C. Weber
Computer Science Department
University of Rochester

Abstract

In this paper, we show how degrees of belief about causal predictions can be derived from statists about the truth of properties over time. By using statistical information analogous to traditional non-statistical rules of causation and knowledge about a specific time point, predictions can be made about both the change and persistence of properties for the next time point with some degree of belief. We show how to incrementally compute this degree of belief by combining statistics conditioned on successively larger subsets of the reasoner's knowledge. Furthermore, we solve the qualification problem through a powerful heuristic that builds these subsets by considering properties with highest impact first. This heuristic ignores relatively unlikely, redundant, or unrelated properties when deriving a prediction, while directing the focus along causal chains. The iterative formula and this heuristic define an algorithm that produces predictions with quickly increasing confidence, allowing computational resources to trade off against accuracy.

1 Introduction

Causal reasoning is the representation and use of the relationships between causes and effects. Traditionally, these relationships have taken the form of implications between sufficient preconditions and necessary future effects, e.g. the following rule about starting your car:

$$\forall t(\texttt{holds}(\texttt{fueled}, t) \wedge \texttt{occurs}(\texttt{turn-key}, t) \rightarrow \texttt{holds}(\texttt{running}, t'))$$

where "t'" is the first time point following "t". The truth of holds(running, 5) follows from the truths of both holds(fueled, 4) and occurs(turn-key, 4). The well-known *qualification problem* [McC77] tells us, however, that in complex realistic domains the antecedent cannot definitively entail the consequent; to do so, it must contain impractically many or even infinitely many preconditions. Thus any causal

prediction must have some degree of uncertainty. The key is to be able to quantify the uncertainty, allowing reasoning to progress in a probabilistic fashion.

Recent work has suggested capturing this relationship between degrees of belief with subjective conditional probability [DK88, Han88]. For example, such a reasoning system might believe the following rule (cf. [DK88]):

$$P(\texttt{holds}(\texttt{running}, t') | \texttt{holds}(\texttt{fueled}, t) \land \texttt{occurs}(\texttt{turn-key}, t)) = \kappa \quad (1)$$

This rule should be read as "κ is the probability that `running` will hold after `turn-key` occurs while `fueled` holds". Despite its intuitive appeal, this approach leaves several important questions:

1. Where do these probabilities come from, i.e. what knowledge about the world does this rule really capture?

2. What is t? Presumably it is a universally quantified variable, but if it is quantified in the probabilisitc metalanguage, what relationship does it have to the temporal objects in the language of properties and events?

3. How does this rule interact with other rules conditioned on different facts? For example, if we also have the rule:

$$P(\texttt{holds}(\texttt{running}, t') | \texttt{holds}(\texttt{cold}, t) \land \texttt{occurs}(\texttt{turn-key}, t)) = \gamma \quad (2)$$

   which assignment of probability is more correct? Which assignment is more practical?

This paper addresses these questions by casting a representation for uncertain causal reasoning within a logic for statistical inference, **Lp** [Bac88b]. Specifically, our answers to the above questions are:

1. The conditionals are statistics from actual domain observations, or perhaps statistical generalizations to which the agent is predisposed.

2. These statistics are part of the object language **Lp** and the variable t stands for a randomization over individuals. An induction mechanism can be used to generate a degree for particular individuals based on relevant statistics.

3. By the popular rule of *specificity* [Bac88b, Eth87, Tou84], conditioning on more knowledge (a more specific *reference class*) leads to more confidence in the degree of belief.[1] Thus a statistic conditioned on both `fueled` and `cold` would

[1] This principle has its limits [Kyb89a]; in particular, it assumes that the sample sizes are arbitrarily large. Otherwise, a very specific statistic may be based on such a small sample size that it loses significance. We discuss this in more detail in [Web89].

be considered more correct than (1) and (2). Since no specificity relationship holds between (1) and (2), they are themselves incomparable. A more practical comparison would decide which statistic's incorporation more profoundly affects the degree of belief of the effect. This leads us to an algorithm that incrementally builds more specific reference classes by considering (heuristically) more important statistics first. Thus the confidence in the prediction depends on the amount of time allowed for the computation. This is a solution to the qualification problem.

The remainder of this paper presents details on these answers. Section 2 describes the statistical logic **Lp**, section 3 introduces the use of statistical statements to capture causal relationships, and section 5 presents an algorithmic approach for solving the qualification problem with these statistical causal theories.

2 A Logic with Statistical Statements

Lp is a superset of standard first-order logic, extended in two basic ways. First, it has a separate domain sort for a field of statistical objects, with defined arithmetic operations. More importantly, it contains an additional syntactic construct for making statistical statements. Given a formula α with a vector of free variables $\vec{x}$, the term $[\alpha]_{\vec{x}}$ evaluates to one of the statistic objects. For example, if $[\mathtt{flys}(x)]_x = .2$, then two out of ten objects in the domain can fly. Such statements are seldom useful when relative to the entire domain, so we conditionalize on other information in the following manner:

$$[\mathtt{flys}(x)|\mathtt{bird}(x)]_x = [\mathtt{flys}(x) \wedge \mathtt{bird}(x)]_x \div [\mathtt{bird}(x)]_x$$

Lp also contains the axioms of probability.

These statements capture useful information about the domain, yet they say nothing about particular objects, e.g. about whether `tweety` can fly. The degrees of belief we seek require an *inductive mechanism* that relates them to the statistics available. The inductive mechanism for **Lp** is as follows: given a ground formula $\alpha(\vec{c})$ containing constants $\vec{c}$ (and no others), the degree of belief in α relative to some knowledge $\rho(\vec{c})$ (written $B(\alpha(\vec{c})|\rho(\vec{c}))$[2]) equals $[\alpha(\vec{x})|\rho(\vec{x})]_{\vec{x}}$. For example:

$$B(\mathtt{flys(tweety)}|\mathtt{bird(tweety)}) = [\mathtt{flys}(x)|\mathtt{bird}(x)]_x$$

Notice that the degree of belief in a fact is always with respect to some knowledge, and different knowledge may inspire different degrees of belief. As is commonly done with statitical inference, we assume that all statements are additionally conditioned on so-called "general knowledge", e.g. the definitions of predicates, taxonomic information, etc. For more information on **Lp** we refer the reader to [Bac88a].

[2]This formula is not a valid term of **Lp** since B is not a function of **Lp**. Thus degrees of belief are outside of the logic, although with some care it could be added.

3 Statistical Causality

Our ultimate goal is to extend a traditional causal reasoning approach so it can deal with uncertainty. We start with a reified temporal logic, where the time points are isomorphic with the integers and therefore discrete ($t' = t+1$). Properties are related to time points via the holds predicate; we adopt the stance that events can be defined in terms of properties, so for simplicity we omit events from our notation. It will be convenient to have the following defintion of property negation:

$$\forall p, t(\mathtt{holds}(\overline{p}, t) \equiv \neg\mathtt{holds}(p, t))$$

In previous work [Web88], we presented a framework for causal reasoning without statistics that used domain-dependent causal rules to prove, when possible, when a property change did or did not take place (thus solving the persistence problem). Here, those causal rules are replaced by conditional statements concerning a particular property ϕ as in the following:

$$[\mathtt{holds}(\phi, t')|\mathtt{holds}(\overline{\phi}, t) \wedge \bigwedge_{i=0}^{\mu} \mathtt{holds}(\psi_i, t)]_{\langle t\rangle} = \mathit{Change}(\phi, \vec{\psi_\mu}) \tag{3}$$

$$[\mathtt{holds}(\phi, t')|\mathtt{holds}(\phi, t) \wedge \bigwedge_{i=0}^{\nu} \mathtt{holds}(\varphi_i, t)]_{\langle t\rangle} = \mathit{Persist}(\phi, \vec{\varphi_\nu}) \tag{4}$$

Rules of type (3) are used to found a degree of belief that ϕ *becomes* true, and rules of type (4) that ϕ *stays* true. Thus together they solve the persistence problem within our statistical setting. Note that there is symmetry in these rules for a property and its negation, i.e.

$$\mathit{Change}(\phi, \vec{\psi_\mu}) = 1 - \mathit{Persist}(\overline{\phi}, \vec{\psi_\mu})$$

so, as would be expected, the behaviour of a property completely determines the behaviour of its negation.

We return to the example of starting a car, and we will represent some simple statistical statements using the facilities defined above. Suppose that the reasoner requires a degree of belief in the fact $\mathtt{holds}(\mathtt{running}, 0')$. This degree will be relative to some knowledge ρ so by the induction mechanism:

$$B(\mathtt{holds}(\mathtt{running}, 0')|\rho(\mathtt{running}, 0)) = [\mathtt{holds}(p, t')|\rho(p, t)]_{\langle p,t\rangle}$$

where ρ is knowledge used to determine the degree of belief. The inductive mechanism replaced the property constant as well as the time point; this feature can lead to the use of statistical generalizations about property types given appropriate knowledge in ρ. However, since in this paper we are more interested in temporal generalizations, we will assume that the general knowledge contains enough information to define the

property in question, thereby establishing the following simplified induction mechanism:

$$B(\texttt{holds}(\texttt{running}, 0')|\rho(0)) = [\texttt{holds}(\texttt{running}, t')|\rho(t)]_{\langle t\rangle}$$

Our choices of ρ will be conjunctions as in (3) and (4). For starting a car, one example might be:

$$[\texttt{holds}(\texttt{running}, t')|\texttt{holds}(\overline{\texttt{running}}, t) \wedge \texttt{holds}(\texttt{ignition}, t)]_{\langle t\rangle} = .9$$

In words, the car starts 90% of the time that its ignition system is engaged. As in (3), we could more succinctly express this as

$$\mathit{Change}(\texttt{running}, \langle\texttt{ignition}\rangle) = .9 \tag{5}$$

Perhaps the car does not like cold weather, signified by the following change in belief when it is also known to be cold:

$$\mathit{Change}(\texttt{running}, \langle\texttt{ignition}, \texttt{cold}\rangle) = .2 \tag{6}$$

This example shows the non-monotonic behaviour of belief; whereas if we based our belief on (5) we might act as if the car will start, if we based our belief on more information in (6) we might act as if it would not. As we add more information to the reference class, we become more confident in the resulting degree of belief. Ideally, we would therefore base belief on the impact of all available knowledge and the problem would be solved. Unfortunately, this approach fails because there is an exponential number of different (very complex) contexts, and the relevant statistic may not be available. Furthermore, even if the statistic is derivable from the knowledge base, the computational effort to do so might be prohibitively expensive. We would like some way to trade reference class specificity for speed, which we do in the next section.

4 Incremental Refinement of Belief

Instead of choosing particular reference classes, our approach will be to start by computing the degree of belief for simple classes, and incrementally refine the result by making the class more complex. In this way, the longer the computation progresses the more confidence the answer will be attributed. If the reasoner decides that it must act, then the current result can be used at the current accuracy. Similarly, if an estimate of the current accuracy is made and found to fall within desired limits, the computation can be stopped. We will do this using *recursive Bayesian updating* [Pea88, p. 37] which is based on the following theorem from the definition of conditionalization:

$$[\texttt{holds}(\phi, t')|\bigwedge_{i=0}^{n} \texttt{holds}(\psi_i, t)]_{\langle t\rangle} = [\texttt{holds}(\phi, t')|\textstyle\bigwedge_{i=0}^{n-1} \texttt{holds}(\psi_i, t)]_{\langle t\rangle} \times \frac{[\texttt{holds}(\psi_n, t)|\bigwedge_{i=0}^{n-1} \texttt{holds}(\psi_i, t) \wedge \texttt{holds}(\phi, t')]_{\langle t\rangle}}{[\texttt{holds}(\psi_n, t)|\bigwedge_{i=0}^{n-1} \texttt{holds}(\psi_i, t)]_{\langle t\rangle}}$$

According to this formula, a property ψ_n that is known to hold may be added to the previous evidence for ϕ holding to produce a statistic for the compound reference class. The impact of the new property is determined by the quotient in the above formula, which we will abbreviate as $Impact(\psi_n, \phi, \vec{\psi}_{n-1})$. When reasoning about change, the base case ψ_0 will be $\overline{\phi}$, and when reasoning about persistence $\psi_0 = \phi$. By the principle of specificity, the reasoner's confidence in the use of the statistic for inducing a particular time's degree of belief should be greater for the new reference class. In this way, each iteration makes the reasoner more confident in the degree of belief.

The *rate* of confidence increase should depend on the order that the property are considered. For example, when starting a car, a statistic based on the weather and condition of the car would inspire more confidence than one based on the arrangement of furniture at the office, at least in realistic scenarios. At each iteration, we employ a heuristic that we call *highest impact first (HIF)*. This heuristic chooses the property whose quotient will most affect the previous degree of belief. Since the new quotients are multiplied in, those with the most impact will be those farthest from the multiplicative idenity, namely 1. This amounts to first considering properties that have a strongly positive or strongly negative influence on the degree of belief in the change. For example, consider the following car-starting scenario involving the following five properties (together with their informal meanings):

r	the car is running
i	the car's ignition is engaged
c	the weather is cold
d	the weather is damp
b	the car has a new battery

For brevity, we will use the shorthand $\psi(t)$ for $\mathtt{holds}(\psi, t)$. We are interested in evaluating whether the car will start at time $0'$ given $\overline{\mathtt{r}}(0)$, $\mathtt{i}(0)$, $\mathtt{c}(0)$, $\mathtt{d}(0)$, and $\mathtt{b}(0)$. The induction mechanism says that our degree of belief in $\mathtt{r}(0')$ with respect to some subset of our knowledge $\rho(0)$ and equals the statistical statement $[\mathtt{r}(t')|\rho(t)]_{\langle t\rangle}$. We will incrementally compute ρ, starting with the base case $[\mathtt{r}(t')|\overline{\mathtt{r}}(t)]_{\langle t\rangle} = \epsilon$. The statistical object ϵ is, of course, very small assuming that the car is started often with respect to all time points. Our method proceeds by computing the impact quotients for each property (since they have no data dependencies, they may be computed in parallel). The impact of the ignition system property i is described by:

$$Impact(\mathtt{i}, \mathtt{r}, \langle \overline{\mathtt{r}} \rangle) = \frac{[\mathtt{i}(t)|\overline{\mathtt{r}}(t) \wedge \mathtt{r}(t')]_{\langle t\rangle}}{[\mathtt{i}(t)|\overline{\mathtt{r}}(t)]_{\langle t\rangle}} = \frac{.99}{1.04\epsilon} = \frac{.95}{\epsilon}$$

Here we supplied some reasonable numbers for this example; the numerator captures that engaging the ignition system is almost certain given that the car starts (leaving some room for hot-wiring), and the denominator captures that most of the time that

the car is idle the ignition system is not engaged. The constant ϵ is used to ensure that although the denominator is probably quite small, it should at least be slightly larger than the small number $[\mathtt{r}(t')|\overline{\mathtt{r}}(t)]_{\langle t\rangle}$. For cold, the impact is the opposite; it is less likely that it is cold given that the car started (the numerator) than it is given that the car is idle (the denominator), giving us an impact less than one, say

$$Impact(\mathtt{c}, \mathtt{r}, \langle\overline{\mathtt{r}}\rangle) = .042/.1 = .42$$

Similarly, we choose .7 as the impact of d and 1.2 as the impact of b. The HIF heuristic is now applied, and it is easily seen that i has the most impact on the degree of belief, since ϵ is very small. Thus the ignition system is factored in, which drastically pulls the degree of belief into the realm of the probable at .95, capturing that most of the time the car starts when the ignition is engaged. Although this belief is large, our confidence in the belief is still low since we have not used much of our knowledge in founding it. The reason for this is demonstrated by examining the impact quotients on the next iteration; given that our other properties are independent of the ignition system (reasonable in this example), the impacts do not change. The impact of i *does* change, since now $\mathtt{i}(t)$ is conditioned on itself, i.e.

$$Impact(\mathtt{i}, \mathtt{r}, \langle\overline{\mathtt{r}}, \mathtt{i}\rangle) = \frac{[\mathtt{i}(t)|\overline{\mathtt{r}}(t) \wedge \mathtt{r}(t') \wedge \mathtt{i}(t)]_{\langle t\rangle}}{[\mathtt{i}(t)|\overline{\mathtt{r}}(t) \wedge \mathtt{i}(t)]_{\langle t\rangle}} = 1$$

This impact will be ignored by the HIF heuristic, thus preventing redundant consideration in a natural way. Instead, HIF will choose c's impact of .42, decreasing our belief to:

$$[\mathtt{r}(t')|\overline{\mathtt{r}}(t) \wedge \mathtt{i}(t) \wedge \mathtt{c}(t)]_{\langle t\rangle} = .95(.42) = .4$$

On the next iteration, the new property c changes the impacts of both d and b, but in two different ways. Since many damp days are also cold days, the knowledge $\mathtt{c}(t)$ in a sense already covers much of $\mathtt{d}(t)$'s statistical impact. Therefore, the impact of d becomes closer to one, say .95. On the other hand, the fact that it is cold *augments* the impact of the new battery (since new batteries perform better than old batteries in cold weather), say raising it from 1.2 to 2. The HIF heuristic chooses this constructive impact, and the new degree of belief is $.4(2) = .8$, a reasonable chance of success. When we consider the next iteration, there are no properties remaining that have significant impact, so (depending on the accuracy desired for the prediction) it may be reasonable to stop augmenting the reference class. The final answer would be:

$$B(\mathtt{r}(0')|\overline{\mathtt{r}}(0) \wedge \mathtt{i}(0) \wedge \mathtt{c}(0) \wedge \mathtt{b}(0)) = Change(\mathtt{r}, \langle\mathtt{i}, \mathtt{c}, \mathtt{b}\rangle) = .8$$

A belief in a persistence could be computed in an analogous way, simply by using the different base case $[\mathtt{r}(t')|\mathtt{r}(t)]_{\langle t\rangle}$.

The computation is even easier to orchestrate if it works with the negative logarithm of the belief $V_t(\alpha) = -\log([\alpha]_{(t)})$, i.e.

$$
\begin{aligned}
V_t(\texttt{holds}(\phi,t)|\textstyle\bigwedge_{i=0}^{n}\texttt{holds}(\psi_i,t)) = \\
&V_t(\texttt{holds}(\phi,t)|\textstyle\bigwedge_{i=0}^{n-1}\texttt{holds}(\psi_i,t))+ \\
&V_t(\texttt{holds}(\psi_n,t)|\textstyle\bigwedge_{i=0}^{n-1}\texttt{holds}(\psi_i,t)\wedge\texttt{holds}(\phi,t))- \\
&V_t(\texttt{holds}(\psi_n,t)|\textstyle\bigwedge_{i=0}^{n-1}\texttt{holds}(\psi_i,t)))
\end{aligned}
$$

This form computes degrees of belief through addition and subtraction of positive reals. The HIF heuristic is also simpler, since it now selects impacts with the largest absolute values, with an impact of 0 being no impact.

We started this section by saying that statistics such as $Change(\phi, \vec{\psi_\mu})$ may not be readily available for all states of knowledge $\vec{\psi_\mu}$. Our remedy was to incrementally build an appropriate statistic, but in the process we required that a large number of other statistics be available to compute the impact quotients. We feel, however, that these latter statistics are more practically accessible. The quotient denominators are statistical relationships between properties *for the same time point*, and as suggested by the above example, they are often independent. The numerators involve the same relationships, but also conditioned on the effect. This relationship can be easy to determine, especially for highly necessary properties such as i (ignition), since it involves assessing the causes given the effect rather than the other way around. Also, since we do not necessarily need to iterate until the entire state of knowledge is incorporated in the reference class, the statistics needed (especially early in the iteration) will be much simpler.

Another possible criticism of our approach is that the knowledge used to form the reference class must be certain, e.g. in the above example we had to *know* $\overline{\texttt{r}}(0) \wedge \texttt{c}(0) \wedge \texttt{i}(0) \wedge \texttt{b}(0)$ with a degree of belief of 1. We can relax this somewhat by only requiring that reference class components be *practical certainties* [Kyb89b], i.e. close to 1. Consider the standard qualification problem example, involving the fact that a potato jammed into a car's tailpipe prevents it from starting. Such circumstances rarely happen, so in practice this fact should be ignored. If spud(t) is the ϵ-unlikely fact that a potato is in the tailpipe at t, then $\overline{\texttt{spud}}(\texttt{t})$ is a practical certainty and can therefore be included in the reference class. Note that the HIF heuristic will actually ignore it, because if we assume that $\overline{\texttt{spud}}$ is independent of our current reference class ρ, we find that:

$$
Impact(\overline{\texttt{spud}}, \texttt{r}, \langle\overline{\texttt{r}}, \rho\rangle) = \frac{[\overline{\texttt{spud}}(t)|\overline{\texttt{r}}(t) \wedge \texttt{r}(t') \wedge \rho(t)]_{(t)}}{[\overline{\texttt{spud}}(t)|\overline{\texttt{r}}(t) \wedge \rho(t)]_{(t)}} = \frac{1}{1-\epsilon} \approx 1
$$

Thus $\overline{\texttt{spud}}$ has very little impact, despite the fact that spud has a devestating impact (the numerator of its quotient is zero). In this way we have solved the qualification problem for very unlikely properties as well as properties with little impact.

5 Summary and Conclusions

In this paper, we have shown how degrees of belief about causal predictions can be derived from statists about the truth of properties over time. By using statistical information analogous to traditional non-statistical rules of causation and knowledge about a specific time point, predictions can be made about both the change and persistence of properties for the next time point with some degree of belief. We have shown how to incrementally compute this degree of belief by combining statistics conditioned on successively larger subsets of the reasoner's knowledge. Furthermore, we solved the qualification problem through a powerful heuristic that builds these subsets by considering properties with highest impact first. This heuristic ignores relatively unlikely, redundant, or unrelated properties when deriving a prediction, while directing the focus along causal chains. The iterative formula and this heuristic define an algorithm that produces predictions with quickly increasing confidence, allowing computational resources to trade off against accuracy.

We are far from done with this line of research. We are currently investigating the following extensions:

- Parallel algorithms where each property is related to a processor that computes the impact of that property on the current prediction. The results are sent to a central arbiter who picks the largest impact and then informs the other processors of the choice, and then the process repeats until the desired confidence is reached.

- An iterative algorithm that computes the belief in a prediction using the belief in a reference class rather than its practical certainty. Note that this is underdetermined, unless the evidence for the belief in the reference class can be applied directly to the belief in the prediction. It does, however, bound the probability due to the following theorem:

$$[\alpha]_{(t)} = \frac{[\alpha|\rho]_{(t)}[\rho]_{(t)}}{[\rho|\alpha]_{(t)}} \geq [\alpha|\rho]_{(t)}[\rho]_{(t)}$$

  Thus the belief can be determined if $[\rho|\alpha]_{(t)}$ is known, or bounded from below if not. An upper bound can be derived analogously by considering $\overline{\alpha}$.

- The relationship to reasoning about knowledge and planning. The HIF heuristic can lead the reasoner to properties that, if known to be true or false, would generate a large impact on the degree of belief. This analysis would inspire the performance of appropriate tests to increase the confidence in the degree of belief. Similarly, information about the utility of different outcomes could be used to inspire actions by the reasoning agent that would increase the future degree of belief in a property with a high utility.

For more details and issues we refer the reader to [Web89].

Acknowledgements

I would like to thank Josh Tenenberg, Henry Kyburg, James Allen, Fahiem Bacchus, Leo Hartman and Len Schubert for enlightening discussion and generous encouragement towards the development of these ideas.

References

[Bac88a] Fahiem Bacchus. *Representing and Reasoning with Probabilistic Knowledge.* PhD thesis, University of Alberta, Fall 1988.

[Bac88b] Fahiem Bacchus. Statistically founded degrees of belief. In *Proceedings of CS/CSI*, pages 59–66, June 1988.

[DK88] Thomas Dean and Keiji Kanazawa. Probabilistic temporal reasoning. In *Proceedings of AAAI-88*, pages 524–528, August 1988.

[Eth87] David W. Etherington. More on inheritance hierarchies with exceptions: Default theories and inferential distance. In *Proceedings of AAAI-87*, pages 352–357, 1987.

[Han88] Steve Hanks. Representing and computing temporally scoped beliefs. In *Proceedings of AAAI-88*, pages 501–505, August 1988.

[Kyb89a] Henry E. Kyburg, Jr. Beyond specificity. Submitted to IJCAI-89, 1989.

[Kyb89b] Henry E. Kyburg, Jr. Evidential and practical certainty. Submitted to IJCAI-89, 1989.

[McC77] John McCarthy. Epistemological problems of artificial intelligence. In *Proceedings of IJCAI-77*, pages 1034–1044, Cambridge, MA, 1977.

[Pea88] Judea Pearl. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* Morgan Kaufman, 1988.

[Tou84] David S. Touretzky. Implicit ordering of defaults in inheritance systems. In *Proceedings of AAAI-84*, pages 322–325, 1984.

[Web88] Jay C. Weber. Solving the persistence problem for an explicit temporal representation. submitted to the First International Conference on Principles of Knowledge Representation and Reasoning, November 1988.

[Web89] Jay C. Weber. *Principles and Algorithms for Causal Reasoning with Uncertain Knowledge.* PhD thesis, University of Rochester Computer Science Department, May 1989.